

UNIVERSIDADE FEDERAL DO PARANÁ

BRUNO VINICIUS FABBRIS COSTA

**AVALIAÇÃO DE POLIMORFISMOS EM POSSÍVEIS SÍTIOS ALVOS DE MIRNAS  
EM LINHAGENS CELULARES TUMORAIS E NORMAL DE MAMA**

CURITIBA

2017

BRUNO VINICIUS FABBRIS COSTA

**AVALIAÇÃO DE POLIMORFISMOS EM POSSÍVEIS SÍTIOS ALVOS DE MIRNAS  
EM LINHAGENS CELULARES TUMORAIS E NORMAL DE MAMA**

Trabalho de Conclusão de Curso I, apresentado ao  
Curso de Graduação de Biomedicina da Universidade  
Federal do Paraná, como requisito parcial à obtenção do  
título de Bacharel em Biomedicina.

Orientadora: Profa. Dra. Daniela Fiori Gradia  
Co-orientador: Dr. Gabriel Adelman Cipolla

CURITIBA

2017

## **AGRADECIMENTO**

Primeiramente eu queria agradecer a minha mãe por aguentar eu surtando por causa desse trabalho, e por me dar conselhos que ajudaram a eu manter a calma

A minha orientadora Profa Dra Daniela Fiori Gradia pelo apoio, ajuda, ensinamentos e pela amizade e ao meu co-orientados Dr Gabrieleal Adelman Cipolla pelas instruções fornecidas.

A Profa Dra Karin Prado Braun por me escolher como aluno e pela oportunidade de realizar esse trabalho.

A Débora pelo carinho e amizade e a Verônica, Luciana e Hellen pela ajuda fornecida na realização dos ensaios.

A toda equipe do LGMH pelo apoio nesses últimos dois anos.

A Prof Dra Maria Luiza Petzl-Erler por disponibilizar o laboratório e os equipamentos necessários para a realização desse trabalho.

As minhas amigas da biomedicina, Valentina, Bruna Priscilla, Ana, Juliana, Gabriela, Rafaela e a Isabela por me darem apoio, força e sempre acreditarem em mim.

## RESUMO

Os polimorfismos de nucleotídeo único (do inglês SNP, *single nucleotide polymorphism*) são o tipo mais comum de variação genética encontrada em populações. Nos genes, esses SNPs são encontrados nas regiões codificantes e nas regiões regulatórias, podendo interferir na regulação da expressão do gene. Os microRNAs (miRNAs) são responsáveis pela regulação pós-transcricional da expressão gênica através de sua ligação à região 3' não traduzida (3'UTR) dos RNA mensageiros, reduzindo a produção proteica. SNPs localizados na região 3' UTR podem interferir na regulação por miRNAs, sendo capazes de fortalecer, impedir ou criar novos sítios de ligação para diferentes miRNAs. SNPs em regiões não codificantes são frequentemente associados a doenças, inclusive ao câncer. Em um trabalho realizado anteriormente a esse, por membros do Laboratório de Genética Molecular Humana foram preditos SNPs localizados nos sítios alvos de miRNAs que, por estudos de associação e de efeito de variantes genéticas sobre a expressão gênica, foram relacionados ao câncer de mama. Para a validação da predição realizada, foram escolhidos os SNPs rs8028 do gene *LMOD1*, e rs13005388 e rs71350312 do gene *STAM2*. O DNA obtido a partir de linhagens celulares tumorais e normal de mama foi utilizado para a realização da reação de sequenciamento pelo Método de Sanger, após amplificação por PCR de parte da região 3'UTR dos genes, onde estão localizados os SNPs de interesse, para determinar quais são os genótipos das diferentes linhagens celulares. Dos três SNPs amplificados, apenas dois foram sequenciados, o rs8028 e o rs71350312. Os resultados obtidos mostraram que existe variação genética entre as linhagens de mama. Para o sequenciamento do rs8028, que tem o alelo C ou T, duas linhagens apresentaram genótipo heterozigoto, duas linhagens foram homozigotas para C, e quatro linhagens homozigotas para T. Já o rs71350312, que também tem os alelos T e C, apresentou uma linhagem heterozigota, uma linhagem homozigota para C, e cinco linhagens homozigotas para T, uma das linhagens não teve seu genótipo determinado. Por fim, o sequenciamento do rs13005388 revelou menos variabilidade genética tendo 6 linhagens homozigotas para o alelo A, uma heterozigota (A e G), e uma linhagem que também não foi possível determinar o genótipo.

Palavras-chave: câncer de mama; *STAM2*; *LMOD1*; polimorfismo.

## ABSTRACT

Single nucleotide polymorphism (SNP) are the most common type of genetic variation detect in population. This SNPs are found inside the genes in coding and regulatory regions, might interfering at gene expression. The microRNAs (miRNAs) are responsible for post-transcriptional regulation of gene expression, through it's binding at 3' untranslated region (3'UTR) at messenger RNA (mRNA) reducing the gene translation. SNPs located in 3'UTR may change the binding affinity between miRNAs and its targets, making this binding stronger, weaker or creating new bindings for others miRNAs. SNPs in non-coding regions are often associated with diseases, including cancer. Recently, our research group develop a work that integrated different levels of data to create predictions of SNPs that interfere with miRNAs target sites, genomic association and effect of genetic variants on gene expression specific to breast cancer. The validation of the predictions occurred using the SNP: rs8028 from *LMOD1* gene and rs71350312 and rs13005388 from *STAM2* gene. The DNA got from cancer and normal cell lines was used to the sequence reaction by Sanger's method, after a PCR amplification of the genes 3'UTR, where the SNPs are located, to determine their genotype. Starting with 3 SNPs, only two were sequenced: rs8028 and rs71350312. The results were: the rs8028 have two alleles C or T, and from all lineages studied, two have the heterozygous genotype, other two lineages are homozygous for C allele and the four left have the homozygous genotype for the allele T. The rs71350312 that also have C and T alleles, shows only one lineage homozygous for C, one heterozygous, five homozygous for T, and one genotype couldn't be determined. The rs13005388 showed six lineages homozygous for A, one heterozygous (A and G), and the last one the genotype couldn't be determined too.

Key words: breast cancer; *STAM2*; *LMOD1*; polymorphism

## LISTA DE FIGURAS

FIGURA 1: SÍNTESE DO MIRNA .....	16
FIGURA 2 VARIAÇÃO NA 3'UTR EM SÍTIO ALVO DE MIRNA.....	17
FIGURA 3: REPRESENTAÇÃO DO ESCRT-0 .....	20
FIGURA 4: ILUSTRAÇÃO DOS DOMÍNIOS DE TMOD1 E LMOD1 .....	22
FIGURA 5: ELETROFORESE DA PCR DA REGIÃO 3'UTR DOS GENES <i>STAM2</i> E <i>LMOD1</i> .....	30
FIGURA 6: GEL DE ELETROFORESE DA LINHAGEM MDA-MB-231 .....	31

## LISTA DE TABELAS

QUADRO 1. CARACTERÍSTICAS DAS LINHAGENS DE MAMA.....	26
QUADRO 2. INICIADORES ESPECÍFICOS PARA REGIÃO CONTENDO O SNP DE INTERESSE. ....	27
QUADRO 3 .CICLOS DA PCR PARA AMPLIFICAÇÃO DA REGIÃO CONTENDO O SNP DE INTERESSE. ....	27
QUADRO 4. CICLOS DA PURIFICAÇÃO ENZIMÁTICA DOS PRODUTOS AMPLIFICADOS POR PCR .....	28
QUADRO 5. ETAPAS DO CICLO DE SEQUENCIAMENTO DAS AMOSTRAS PURIFICADAS.....	28
QUADRO 6. GENÓTIPO DAS LINHAGENS DE MAMA .....	31
QUADRO 7. PREDIÇÕES REALIZADAS PARA OS MIRSNP'S .....	33
QUADRO 8. TARGET SCAN HUMAN .....	35

## SUMÁRIO

<b>1 INTRODUÇÃO</b>	9
<b>2 OBJETIVOS</b>	11
2.1 OBJETIVO GERAL	11
2.2 OBJETIVO ESPECÍFICO	11
<b>3. REVISÃO BIBLIOGRÁFICA</b>	12
3.1 CÂNCER E CÂNCER DE MAMA	12
3.2 CLASSIFICAÇÕES MOLECULARES DOS TUMORES DE MAMA	13
3.3 SUSCEPTIBILIDADE GENÉTICA AO CÂNCER DE MAMA	14
3.4 <i>STAM2</i>	17
3.5 <i>LMOD1</i>	21
<b>4. MATERIAL E MÉTODOS</b>	24
4.1 OBTENÇÃO DOS SNPS CANDIDATOS	24
4.1.1 INTEGRAÇÃO DE DADOS E CLASSIFICAÇÃO BAYESIANA	24
4.1.2 INTEGRAÇÃO COM DADOS DE GWAS	24
4.1.3 INTEGRAÇÃO DE DADOS DE eQTL	25
4.2 LINHAGENS UTILIZADAS	25
4.3 GENOTIPAGEM	26
<b>5 RESULTADOS</b>	30
5.1 AMPLIFICAÇÃO DA REGIÃO 3'UTR DO GENES <i>STAM2</i> E <i>LMOD1</i>	30
5.2 SEQUENCIAMENTO	31
5.3. INTERAÇÃO ENTRE MIRNAS PREDITOS E O SNP	32
<b>6. DISCUSSÃO</b>	36
<b>7. CONCLUSÃO</b>	39
<b>8. PERSPECTIVA</b>	40
<b>9 REFERÊNCIAS</b>	41



## 1 INTRODUÇÃO

O desenvolvimento biológico e as propriedades fenotípicas que um tumor apresentar serão determinadas pelas variações genéticas encontradas nele. RNAs não codificantes, como os microRNAs (miRNAs), têm o potencial de influenciar essas características através de regulação pós-transcricional dependente de pareamento no formato Watson e Crick. Eles podem interferir em alguns mecanismos muito importantes para o desenvolvimento tumoral, como na autossuficiência da sinalização do crescimento, na evasão da apoptose, no potencial de replicação ilimitado, na angiogênese, na metástase, entre outros. Assim, esses miRNAs apresentam um potencial de serem utilizados como biomarcadores para o diagnóstico do câncer, ou até como alvos terapêuticos (BERINDAN-NEAGOE *et al.*, 2014).

Os miRNAs conseguem parear com a região 3'UTR dos RNAs mensageiros (RNAm) e, por essa razão, variações genéticas, como os polimorfismos de nucleotídeo único (SNP) nessa região, podem interferir com a afinidade do miRNA ao seu sítio alvo; essas mudanças nos pares de base podem fortalecer, inibir ou criar novos sítios de ligação para os miRNAs (BARTEL, 2004).

Foi desenvolvido um projeto para identificar SNPs em sítios alvos de miRNAs que pudessem estar envolvidos na susceptibilidade ao câncer de mama. Para isso, foram utilizadas três ferramentas de bioinformática (PolymirSNPTS, miRSNPscore e o MirSNP), cada uma com um algoritmo próprio para prever o efeito de SNPs na ligação de miRNAs (MirSNPs). Associado a isso, a consulta a bancos de dados públicos forneceu informações de SNPs relevantes no contexto do câncer de mama: SNPs associados ou que marcam a expressão de genes envolvidos na patogênese da doença. Para reunir a informação fornecida pelas três ferramentas e chegar a um conjunto único de MirSNPs, foi utilizado a Estatística Bayesiana Ingênua. As informações obtidas foram integradas com dados de estudos genômicos de associação (GWAS, do inglês *genome-wide association studies*) e de traços fenotípicos com base genética poligênica (eQTL, do inglês *expression quantitative trait loci*), para assim identificar MirSNPs com capacidade de afetar a expressão de seus genes e que estão associados ao câncer de mama.

Assim, um conjunto de MirSNPs, localizados em diferentes genes foram selecionados. Contudo para a realização deste trabalho, apenas três MirSNPs foram

escolhidos para realizar a validação funcional do trabalho de bioinformática. Os SNPs são: rs8028 do gene *LMOD1*, e rs13005388 e rs71350312 do gene *STAM2*. Tais SNPs foram sequenciados em oito linhagens diferentes de câncer de mama.

## 2 OBJETIVOS

### 2.1 OBJETIVO GERAL

Validar a estratégia de busca por MirSNPs originalmente baseada na: (1) unificação de ferramentas de bioinformática de predição de alvos de miRNAs no contexto de variação genética em sítios-alvo; e (2) na consulta a bancos de dados públicos sobre associação genética e diferença de expressão no câncer de mama.

### 2.2 OBJETIVO ESPECÍFICO

Determinar os genótipos de linhagens celulares normais e tumorais de mama para os SNPs rs8028 do gene *LMOD1* e rs13005377 e rs71031288 do gene *STAM2*.

### 3. REVISÃO BIBLIOGRÁFICA

#### 3.1 CÂNCER E CÂNCER DE MAMA

O corpo humano é formado por células que naturalmente crescem e se dividem. Quando ficam senis, ou perdem sua função, ocorre o processo de morte celular e novas células as substituem. Porém, existem momentos em que o controle desse crescimento é perdido, decorrente de mutações em genes específicos, fazendo essas células se multiplicarem desordenadamente, promovendo o câncer. O nome dado a essa patologia corresponde a uma coleção de doenças relacionadas pelo fato delas poderem ocorrer em qualquer parte do corpo, em qualquer tipo celular. Tais condições são classificadas como benignas, quando ficam confinadas ao tecido de origem, ou malignas, quando se espalham para tecidos circundantes ou até mesmo para regiões distantes daquela que é considerada primária (NATIONAL CANCER INSTITUTE, 2017).

O câncer de mama é uma doença complexa, com alta variabilidade genética, fenotípica e comportamental, e ocorre quando existe um crescimento anormal das células do tecido mamário. As células tumorais provenientes dos ductos ou dos lóbulos mamários correspondem à maior parte dos cânceres de mama maligno denominados carcinomas. Em ocasiões mais raras, o câncer de mama pode ser gerado de outros tipos celulares, como das células estromais, ou dos vasos linfáticos e sanguíneos, e por essa razão recebem o nome de sarcomas e linfomas. (AMERICAN SOCIETY OF CANCER, 2017).

Na classificação das neoplasias que mais acometem os seres humanos, o câncer de mama aparece em segundo lugar. Em 2012, foram diagnosticados aproximadamente 1,7 milhão de novos casos (cerca de 11,9% de todos os casos de câncer no mundo). Em relação a mortalidade, aparece em quinto lugar com 522 mil mortes causadas pela doença (FERLAY *et al.*, 2014).

No Brasil o câncer de mama é o segundo mais comum nas mulheres, perdendo apenas para o câncer de pele não-melanoma, e é responsável por 28% de novos casos de câncer a cada ano. A estimativa da incidência de novos casos é de 57.960 mil pessoas desenvolvendo a doença no país entre 2016 e 2017. Em 2013, foi registrado um número de mortes de 14.206 mulheres em função do câncer de

mama, o que correspondeu a 2,72% das causas de óbito de mulheres no país (INCA, 2017).

### 3.2 CLASSIFICAÇÃO MOLECULAR DOS TUMORES DE MAMA

Perou *et al* (2000) desenvolveram um trabalho que procurou classificar os diferentes tipos de carcinomas de mama através das marcações moleculares que eles apresentavam. Para isso, utilizaram o perfil de expressão gênica global das células de tumores de mama, uma vez que, dependendo do conjunto de genes que estão sendo expressos, as células vão apresentar diferentes taxas de crescimento e utilizar vias de sinalização específicas que irão determinar suas características. Com isso, diferenciaram as linhagens em duas classes, aqueles que possuem os receptores hormonais (RH) – o receptor de estrogênio (RE) e o receptor de progesterona (RP) – os quais eles classificaram em *luminal-like*, e as linhagens que não possuíam esses RH. Dentro desta última classe existem os subgrupos: HER2<sup>+</sup>, que possuem o receptor do fator de crescimento epidermal 2 de humanos (HER2) superexpresso e *basal-like*, células que têm a expressão semelhante às das células epiteliais basais (Mioepitelial).

Entre os adenocarcinomas de mama, temos aqueles que expressam os RH, que são classificados no tipo *luminal-like*, e que se dividem em luminal A e B. Por marcação imunohistoquímica, ambos apresentam os RE e/ou RP, entretanto células pertencentes ao subtipo luminal A não apresentam a superexpressão da proteína HER2 (TELANG *et al.*, 2017), enquanto as células do subtipo luminal B podem ou não ter esse receptor superexpresso (CIRQUEIRA *et al.*, 2011). Mas a principal diferença entre esses dois subtipos é a taxa de proliferação celular que no grupo luminal B é mais elevada do que no luminal A, conferindo aos pacientes com esse subtipo de câncer de mama um pior prognóstico (SORLIE *et al.*, 2001). O subtipo luminal A é o mais encontrado, e também o que melhor responde aos tratamentos hormonais sistêmicos, como a utilização de antagonistas hormonais (tamoxifeno) e inibidores da aromatase (CHRYSANTHOU *et al.*, 2017; WEIGEL; DOWSETT, 2010).

A proteína HER2 (CD34) é uma tirosina quinase que faz parte da família dos receptores dos fatores de crescimentos epidermais (HER), e, para ser fosforilada e iniciar a cascata de ativação, precisa formar um heterodímero com outra proteína da família HER. Assim, muitos dos diagnósticos reportados de câncer de mama

pertencem ao subtipo HER2<sup>+</sup>. Essas células superexpressam o receptor HER2, principalmente pela amplificação do número de cópias desse gene no DNA, que, quando ativado, envia sinais ao núcleo para a expressão de genes responsáveis pelo crescimento e divisão celular. Além disso, esse subtipo possui a expressão dos RH diminuída e tem também grande capacidade de metástase. Devido à presença do HER2 na superfície celular, existe a possibilidade de tratamentos, principalmente com anticorpos monoclonais que promovem sua inativação (NAMI; WANG, 2017; KOUTRAS; EVANS, 2008).

O subgrupo dos cânceres *basal-like* apresenta uma rápida divisão mitótica e, além de possuir um padrão de expressão de genes similar ao de células mioepiteliais, possuem características próprias que definem esse subtipo. Entre essas características há a presença de marcadores exclusivos, como o receptor do fator de crescimento epidermal (EGFR), mutações no gene *BRCA1*, e até mesmo em alguns casos a expressão dos RH. Dentro desse subgrupo encontram-se os tumores de mama triplo negativo (TMTN), que não apresentam os receptores RE, RP e HER2. Esses conjuntos de tumores chamam a atenção pelo fácil diagnóstico (a ausência dos receptores é observada em um exame de imunohistoquímica), a alta agressividade e a dificuldade de tratamento, uma vez que inibidores da aromatase, antagonistas hormonais e fármacos que atuam sobre o HER2 não têm efeito. Além disso, eles são responsáveis por 12-24% de todos os cânceres de mama encontrados (RAKHA; ELLIS, 2009).

### 3.3 SUSCEPTIBILIDADE GENÉTICA AO CÂNCER DE MAMA

Por se tratar de uma doença complexa, o câncer tem seu aparecimento e desenvolvimento influenciados por fatores genéticos e ambientais (BOHON, 2017). Entre os fatores de risco que podem aumentar a suscetibilidade das mulheres para o desenvolvimento do câncer de mama existem: a utilização de remédios hormonais, falta de exercício, o consumo de álcool, ganho de peso após a menopausa, nuliparidade, ter engravidado em um período mais tardio, menarca precoce e menopausa tardia (RODGERS *et al.*, 2017).

Já entre os fatores genéticos, a presença de mutações específicas que são herdadas aumenta a chance de desenvolver o câncer. Aproximadamente, cerca de 5 a 10% dos casos de câncer de mama ocorrem pela mutação em dois genes, o

*BRCA1* e o *BRCA2*, sendo necessária a mutação em apenas um deles, para ter um risco mais elevado de ter a doença. Mulheres após os 70 anos tem um risco de apresentar o câncer, entre 40-87%, se portadoras do gene *BRCA1* mutado, e de 27-84% se portadoras do gene *BRCA2* mutado (JARA *et al.*, 2017). Contudo, nem todos os casos hereditários de câncer são atribuídos as mutações nesses genes, assim mulheres com um histórico familiar positivo para câncer de mama, já tem um risco mais elevado de desenvolver a doença (SAKORAFAS, 2003).

Esses fatores genéticos que conferem maior, ou menor, suscetibilidade a doença existem devido a variabilidade genética presente em cada indivíduo. Sendo uma das variações mais comuns os polimorfismos de nucleotídeo único (SNP), nos quais ocorre uma mudança em apenas um par de bases na sequência do DNA (FURRER *et al.*, 2016; COWPER-SALLARI *et al.*, 2012). Muitos desses SNPs (*loci*) estão em regiões reguladoras, não alterando a sequência de aminoácidos da proteína codificada, mas modificando sua expressão (FRAZER *et al.*, 2009).

A 3'UTR (região 3' não traduzida) dos genes é uma dessas regiões regulatórias que, quando o gene é transcrito em um RNA mensageiro maduro (RNAm), tais regiões podem ser alvo de microRNAs (miRNAs) (BORCHERT; LANIER; DAVIDSON, 2006)

Os miRNAs estão presentes em genes espalhados ao longo do genoma, podendo estar localizados em regiões intergenicas, em introns e em exons. Essas moléculas são transcritas pela RNA polimerase II ou, menos regularmente pela RNA polimerase III, na sua forma precursora de pri-miRNA, com centenas de nucleotídeos, apresentando um CAP na extremidade 5' e uma cauda poli-A na extremidade 3'. No núcleo, ribonucleases específicas chamadas DROSHA junto ao cofator DGCR8, clivam o pri-miRNA gerando o pré-miRNA de tamanho variado, entre 60 a 70 nucleotídeos, que apresenta uma conformação em grampo, fazendo esse RNA ter uma estrutura de dupla-fita (FANG *et al.*, 2017).

Esse pré-miRNA, localizado no núcleo vai ser transportado para o citoplasma, com o auxílio da proteína exportina 5, dependente de RAN-GTP. No citoplasma vai haver a enzima DICER, uma RNase que consegue interagir com o pré-miRNA, devido a um domínio dessa enzima que reconhece RNA dupla-fita (dsRNA). Assim ela vai clivar o pré-miRNA em pequenos dsRNA de aproximadamente 22 nucleotídeos (VISONI; CROCE, 2009).

Proteínas da família das Argonautas (AGO), também encontradas no citoplasma, vão reconhecer esse pequeno dsRNA, se ligar a ele e separar as duas fitas, uma vai ser selecionada e a outra descartada. Junto com a AGO, outras proteínas formam um complexo proteico denominado RISC (complexo de silenciamento induzido por RNA) (HAJARNIS *et al.*, 2017). O RISC junto ao miRNA interage com moléculas de RNAm na região 3'UTR, devido à complementaridade das sequências, principalmente da região *seed*, uma região localizada entre o 2º ao 8º nucleotídeo, altamente conservada nos miRNAs. Essa interação pode impedir a tradução ou promover a degradação do transcrito. Em ambos os casos tem-se uma regulação negativa do gene alvo (BORCHERT; LANIER; DAVIDSON, 2006; BARTEL, 2004). (FIGURA 1).

**FIGURA 1: SÍNTESE DO MIRNA**

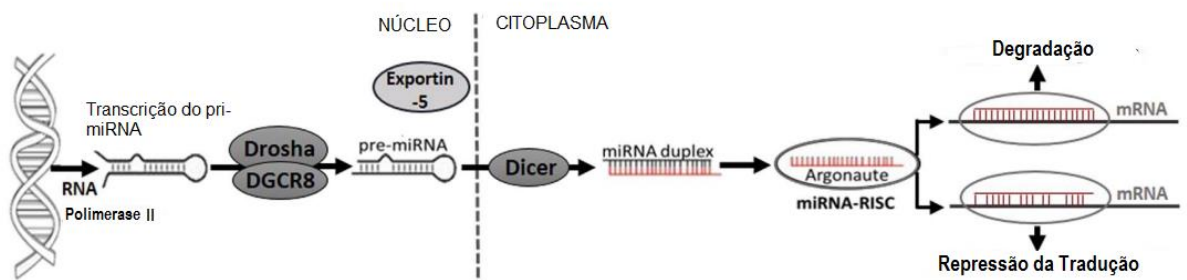


Imagem mostrando o processo de síntese do miRNA maduro. A transcrição do gene pela enzima RNA Polimerase II, em pri-miRNA. A clivagem pela DROCHA/DGCR8, gerando o pré-miRNA, que é transportado para o citoplasma pela Exportina 5. No citoplasma, é clivado novamente por outra enzima, a DICER, formando um miRNA de dupla-fita, que é reconhecido pela proteína Argonauta, e que vai selecionar uma das fitas (nesse caso a de cor vermelha), para fazer parte do RISC. Essa ribonucleoproteína vai interagir com o mRNA alvo gerando a degradação dele, ou reprimindo a sua tradução. Adaptado de BECKETTA; VEYSEYB; LUCOCK (2017).

Assim, polimorfismos na 3'UTR do mRNA podem afetar a relação com o miRNA (FIGURA2). Essas mudanças podem inibir a ligação do miRNA naquele sítio, contudo pode também aumentar a interação entre as duas moléculas, ou ainda criar novos sítios de ligação para diferentes miRNAs. Assim, conhecer o papel que SNPs na 3'UTR desempenham em doenças complexas como o câncer se torna



importante, visto que já foram descritos SNPs associados ao risco de desenvolver as doenças (PRESKILL; WEIDHAAS, 2013).

Neste trabalho foram estudados SNPs encontrados na região 3'UTR de dois genes: *STAM2* e *LMOD1*.

**FIGURA 2** VARIAÇÃO NA 3'UTR EM SÍTIO ALVO DE MIRNA

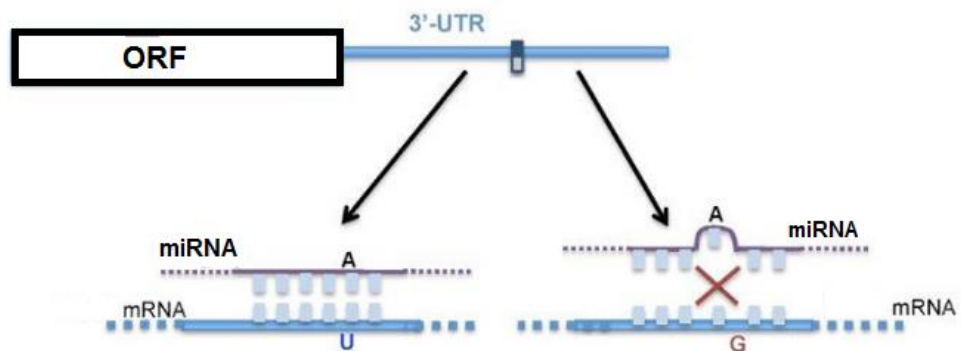


Imagem mostrando que a mudança de um nucleotídeo na região 3'UTR de um mRNA, pode modificar a interação entre ela e o miRNA, nesse caso impedindo a ligação do miRNA. Fonte: Adaptado de PRESKILL; WEIDHAAS (2013).

### 3.4 *STAM2*

O gene *STAM2* está localizado no cromossomo 2q23.3, na fita reversa. Tem 14 exons e codifica para a molécula adaptadora da tradução de sinal 2 (STAM 2). Essa molécula faz parte da família de proteínas STAM, junto da proteína STAM 1, da qual é considerada parálogo. Essas moléculas fazem parte de um complexo proteico denominado ESCRT-0, junto com outra proteína denominada Hrs (GENECARD, 2017).

Já foram descritas algumas funções para a proteína STAM2, como o envolvimento na via de sinalização intracelular gerada pela citocina IL-2. Também, foi relatada a participação dela no desenvolvimento e crescimento de linfócitos T. Mas a principal função descrita foi no processo de degradação de receptores e proteínas de adesão presentes na superfície celular, marcados com ubiquitina,

através da formação dos endossomas (BACHE *et al.*, 2003; TAKATA *et al.*, 2000; ENDO *et al.*, 2000).

A formação dos endossomas ocorre durante o processo de endocitose, no qual diversas vesículas, provenientes da membrana plasmática (MP), vão se fundir no citoplasma formando o endossoma precoce, no qual vão estar presentes as proteínas encontradas na MP. Após algumas etapas de processamento, essa organela é denominada endossoma tardio, e as proteínas encontradas nela podem ter três destinos: retornar à MP, serem enviadas ao complexo de Golgi; ou serem degradadas no lisossoma. Para serem degradadas, os endossomas tardios promovem a formação de vesículas para dentro do seu próprio lúmen, sendo então denominados de corpos multivesiculares (CMV), e as vesículas presentes em seu interior, de vesículas intraluminais (VIL). As proteínas de membrana que serão degradadas vão estar nessas VIL, pois quando os CMV se unirem aos lisossomas vão promover o encontro das VIL com as enzimas de degradação no interior dos lisossomas (ELKIN; LAKODUK; SCHMID, 2016).

O complexo protéico ESCRT (do inglês, *endosomal sorting process required for transport*) foi inicialmente descoberto em 2001 (KATZMANN; BABST; EMR, 2001), e teve como primeira função descrita a participação na formação dos CMV. As proteínas de membrana ubiquitinadas, localizadas na superfície do endossomo, vão ser reconhecidas pelo ESCRT e direcionadas para dentro das VIL (HANSON; CASHIKAR, 2012).

Após a descoberta do ESCRT, que ficou conhecido com ESCRT-I, outros três complexos formados por diferentes proteínas, nomeados como ESCRT-III, ESCRT-II e ESCRT-0, foram descobertos (HANSON, SHIM & MERRILL, 2009). Assim, com mais estudos, outras funções foram sugeridas para esses complexos proteicos, por exemplo, a utilização do ESCRT por vírus envelopados na etapa de liberação do vírus por brotamento na membrana plasmática, para a formação de novos agentes infectantes. Também a participação deles na divisão celular, mais especificamente, no processo de citocinese, no qual a membrana que une as células é separada originando duas células-filhas (HURLEY, 2015).

A proteína ESCRT-0, em mamíferos, é formada por um heterodímero composto pelas proteínas Hrs, com a proteína STAM 1 ou 2. Esse complexo se localiza na membrana endossomal, e tem como função determinar o destino que as proteínas marcadas com ubiquitina vão adquirir (BACHE *et al.*, 2003). A STAM 2

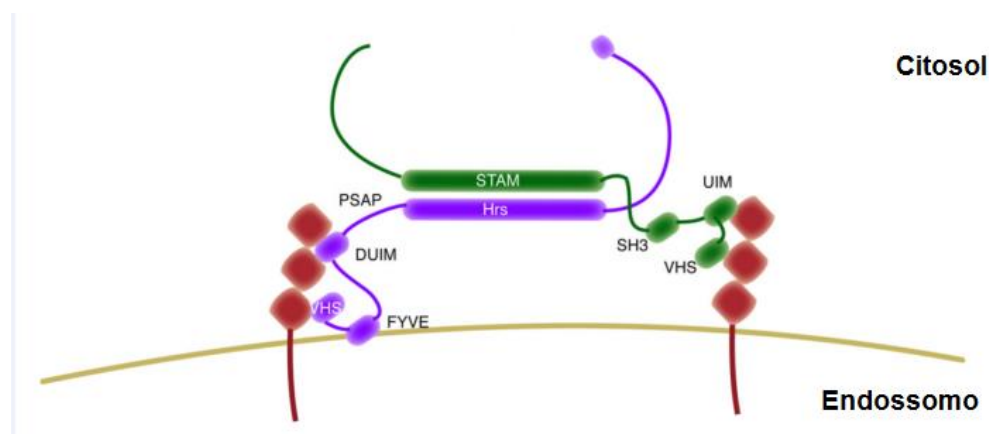
possui 523 aminoácidos e se divide em 5 domínios: o domínio VSH (Vsp27, Hrs, Stam), o motivo de interação a ubiquitina (MIU), o domínio homólogo do Src 3 (SH3), Coiled Coil (CC) e o motivo de ativação de imuno-receptor baseado em tirosina (ITAM) (KAPURALIN *et al.*, 2015).

A proteína Hrs (do ingles, *hepatocyte growth factor-regulated tyrosine kinase substrate*) possui um domínio FYVE que é responsável por permitir que o complexo se associe à membrana dos endossomos precoces, através da sua ligação com a molécula de fosfatidilinositol-trifosfato (PI-3P). A interação que ocorre entre Hrs e proteínas STAM, fazendo elas permanecerem unidas no complexo proteico, se dá pelo domínio CC encontrado em ambas (TAKATA *et al.*, 2000).

Outros domínios encontrados em ambas as proteínas são MIU e o domínio VHS, na região N-terminal, que permitem o complexo proteico interagir com moléculas de ubiquitina. Tais domínios são importantes, pois interagem com as proteínas de membrana marcadas, selecionando o destino delas. Esse destino pode ser: a degradação no lisossomo (como acontece com o receptor do fator de crescimento epidermal- EGFR), ou a reciclagem delas, retornando para a MP (como receptores beta-adrenérgicos) (REN; HURLEY, 2010; HANYALOGU; MCCULLAGH; VON ZASTROW, 2005).

As proteínas da família STAM, ainda podem estar envolvidas com esse processo de outra forma, através do domínio SH3, que pode se relacionar com proteínas desubiquinases, assim controlando essa seleção (HURLEY, 2010). Por fim, as proteínas STAM também tem um domínio ITAM que aparentemente está envolvido na diminuição da expressão das proteínas Jak 2 e Jak 3, pertencentes à família das tirosina-quinase (ENDO *et al.*, 2000) (FIGURA 3).

O ESCRT-0 sozinho não é o responsável pela formação dos CMV, ele precisa interagir como outras proteínas que vão auxiliá-lo nesse processo, e uma delas é o complexo ESCRT-I que tem um importante papel na separação das membranas de dois compartimentos celulares diferentes. Essa interação ocorre através da Hrs que possui o domínio PSAP e interage com o TSG101, proteína que faz parte do ESCRT-I (HANSON; CASHIKAR, 2012; RAZI; FUTTER, 2006).

**FIGURA 3: REPRESENTAÇÃO DO ESCRT-0**

As proteínas que formam o complexo são: Hrs, em roxo, STAM (STAM1 ou STAM2), em verde, e as proteínas de membrana, em vermelho. Também ilustrados os domínios de cada proteína: VSH, UIM, DUIM interagem com a(s) molécula(s) de ubiquitina(s), o FYVE que se associa ao PI-3P na membrana do endossoma, SH3, que interagem com enzimas deubitiquinases, *Coiled Coil*, que são as regiões responsáveis por formar o heterodímero, e o PSAP, que permite a interação com o ESCRT-I. Adaptado de HURLEY (2010).

A relação entre o gene *STAM2* e as demais proteínas da família do ESCRT com o câncer foi discutida por Mattissek e Teis (2014). Nessa revisão, foi comentado o papel que o complexo ESCRT-0 poderia desempenhar como um supressor de tumor. Em um dos experimentos com células de mamíferos, mostrou-se que o aumento da proteína Hrs, diminuiu a presença do EGFR, enquanto a perda da função de ESCRT-0, leva ao efeito oposto, um aumento do EGFR na membrana. Contudo, nas células HeLa, a diminuição de Hrs utilizando siRNA promoveu uma diminuição da capacidade metastática delas, por aumentar a presença de proteínas de adesão celular como a E-Caderina.

Foram realizados estudos que promoveram o nocaute do gene *HRS* e dos genes *STAM1* e *STAM2*, resultando em letalidade embrionária em camundongos (YAMADA *et al.*, 2002; KOMADA & SORIANO, 1999). No estudo do nocaute realizado por Yamada *et al* (2002), a eliminação dos dois genes *STAM* levou a letalidade embrionária, enquanto que a deleção de apenas um deles resultou no desenvolvimento normal do camundongo, indicando que as duas proteínas têm efeitos redundantes no complexo ESCRT-0.

Os SNPs selecionados para esse gene foram dois: o rs13005388 e o rs71350312. O rs13005388 tem dois alelos: A e G, sendo A o alelo ancestral, e G o alelo de menor frequência, com MAF (*minor allele frequency*) de 0,04. O SNP rs71350312 tem os alelos T e C, sendo T o alelo ancestral. A frequência alélica mundial dos dois alelos, de ambos os genes é para os ancestrais de 0,963, e para o alelo derivado de 0,037, tanto que o  $D'=1$  (ENSEMBL, 2017; DLLINK, 2017).

### 3.5 *LMOD1*

O gene da *LMOD1* também conhecido como leiomodina 1, faz parte da família de proteínas tropomodulinas (TMOD), importantes na manutenção dos filamentos de actina. Foi incorporado a essa família em 2001, quando foi visto que regiões da proteína *LMOD1* eram muito semelhantes com domínios das proteínas TMOD, principalmente com a TMOD1. Esse gene é encontrado na região 1q32 e possui 3 exons. Além dele, existem dois genes parálogos, um no cromossomo sete, conhecido como *LMOD2* (CONLEY *et al.*, 2001), e outro no cromossomo três, denominado *LMOD3* (YUEN *et al.*, 2015).

Estudos mostraram que esse gene está mais expresso em células da musculatura lisa (CML), por isso recebeu o prefixo 'leio' em seu nome, que significa liso, em grego. Assim, foi visto que, em camundongos, os tecidos vasculares e viscerais ricos em musculatura lisa, como a aorta, a bexiga, o intestino e o útero, têm maiores níveis de expressão de *LMOD1*, do que órgãos como o coração, o cérebro e a musculatura estriada. Em compensação, os níveis de *LMOD2* e *LMOD3* são altos principalmente na musculatura cardíaca e estriada (YUEN *et al.*, 2015; NANDA; MIANO, 2011; CONLEY, 2001).

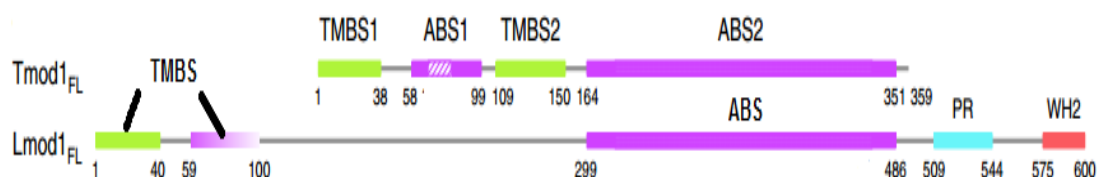
As CML adquirem características fenotípicas próprias durante seu processo de diferenciação, isso ocorre pela ativação de vários genes, em especial por um grupo específico deles que possui a expressão restrita a esse tipo celular. Os genes pertencentes a esse grupo apresentam, na região promotora, o sítio CArG, extremamente conservado nos mamíferos. Neste sítio de 10 pares de bases, se liga um fator de transcrição chamado fator de resposta ao soro (FRS), que, por ser um fator de transcrição fraco, precisa se juntar a outras proteínas, entre elas a miocardina, para promover a expressão desses genes. Foi visto que *LMOD1* tem esse sítio na região promotora e sua expressão ocorre pela ativação dos fatores de transcrição citados, mostrando que esse gene faz parte do grupo que tem expressão

restrita às CML, e por isso pode ter um papel importante no desenvolvimento dessas células (NANDA; MIANO, 2011; MIANO, 2003).

As proteínas da família das tropomodulinas têm como principal função garantir a estabilidade dos filamentos de actina encontrados nas células. Esses filamentos apresentam duas extremidades, e as TMODs são encontradas na extremidade negativa, também chamada de pontiaguda, impedindo a elongação e a despolarização dos filamentos. Outras moléculas também são importantes para garantir a estabilidade dos filamentos de actina, como a tropomiosina (TM), sendo que o TMOD também consegue se associar a ela (CONLEY *et al.*, 2001).

Na porção N-terminal de TMOD se encontram três principais domínios, dois deles ligadores de TM (chamados de TMBS1 e 2) e um domínio de ligação à actina (ABS1), enquanto que a extremidade C-terminal é composta, quase que inteiramente, pelo segundo domínio de ligação à actina (ABS2), fazendo a proteína ter um peso molecular de 40kDA. LMOD1, em contrapartida, tem 64kDA de peso molecular, e é separado em três principais regiões. Na porção N-terminal se encontra a região ligadora de TM, (KOSTYUKOVA, 2007) e na parte central está o domínio de ligação à actina (ABS). Por causa dessas duas regiões, que são semelhantes às encontradas em TMOD, os genes são classificados na mesma família. Além disso, LMOD1 apresenta na sua porção C-terminal um domínio rico em prolina (PR) e um domínio chamado WH2 que se associa com actina. Apesar de terem regiões semelhantes às proteínas TMOD, LMOD1 apresenta funções muito diferentes. Foi visto que LMOD1 está envolvido no processo inicial de formação dos filamentos de actina, processo esse chamado de nucleação, e que os domínios ABS e WH2 desempenham um papel fundamental neste processo (BOCZKOWSKA *et al.*, 2015) (FIGURA 4).

**FIGURA 4:** ILUSTRAÇÃO DOS DOMÍNIOS DE TMOD1 E LMOD1



TMBS1 e 2: sítio de ligação à tropomiosina 1 e 2 de TMOD. ABS1 e 2: sítio de ligação à actina 1 e 2 de TMOD. TMBS: sítio de ligação à tropomiosina do LMOD1. ABS: sítio de ligação à actina de LMOD1. PR: domínio rico em prolina. WH2: região de ligação à actina. Adaptado de BOCZKOWSKA (2015).

Em um estudo publicado por Halim *et al.* (2017), foi relatado o nascimento de uma criança, a partir de um casamento consanguíneo, com a Síndrome de Megabexiga, Microcolon e Hipoperistaltismo Intestinal (SMMHI), uma doença de caráter genético e muito rara que causa aumento da bexiga, gerando uma hidronefrose, diminuição do peristaltismo no intestino e a formação de um microcolon, levando os pacientes a óbito. Essa criança não apresentava nenhuma mutação nos genes já descritos relacionados a síndrome, mas apresentou-se homozigota para uma mutação sem sentido no gene *LMOD1*. Estudos que inativaram o gene *Lmod1* em camundongos revelaram que esses animais desenvolveram os sinais clínicos da SMMHI. Assim, foi comprovado que a ausência de *LMOD1* está relacionada a essa patologia.

Em relação ao câncer, foram vistos casos de neoplasias, nos quais houve perda de heterozigossidade para genes supressores de tumor, ou seja, um dos alelos desses genes foi perdido, com isso restando apenas uma cópia desse gene, que caso apresentasse alguma mudança, podia resultar na perda da função desse gene supressor, o que seria vantajoso para a formação do câncer. Em um estudo realizado com células de câncer de mama, foi visto que em 60% dos casos existia a deleção no *loci* 1q32, em um dos cromossomos, o que gerava a perda da heterozigossidade dos genes nessa região, indicando que nela pode existir um gene supressor de tumor (BENITEZ *et al.*, 1997). A perda da heterozigossidade nesta região também foi descrita em outros tipos de câncer, como no carcinoma colorretal (ZHOU *et al.*, 2008).

O SNP selecionado foi o rs8028, encontrado na 3'UTR de *LMOD1*. Os alelos encontrados para esse SNP são *T* ou *C*, sendo *C* o alelo ancestral, e o MAF de 0.22, para o alelo *T*. Na população europeia, a frequência do alelo *C* é de 0,698, enquanto no mundo, o alelo *C* tem a frequência aumentada para 0,784 e (ENSEMBL, 2017).

## 4. MATERIAL E MÉTODOS

### 4.1 OBTENÇÃO DOS SNPS CANDIDATOS

#### 4.1.1 INTEGRAÇÃO DE DADOS E CLASSIFICAÇÃO BAYESIANA

Tendo em vista o objetivo de identificar SNPs localizados na 3'UTR de genes de interesse e que possam estar em sítios-alvo de miRNAs (MirSNPs), um trabalho foi realizado anteriormente a esse, por diferentes membros do LGMH, sendo que neste projeto, foram utilizados três bancos de dados: PolymiRTS Database v.3 (<http://compbio.uthsc.edu/miRSNP/>) (BHATTACHARYA, ZIEBARTH & CUI, 2014), miRSNP (<http://202.38.126.151/hmdd/mirsnp/search/>) (LIU *et al.*, 2012); e miRSNPscore (<http://www.bigr.medisin.ntnu.no/mirsnpscore/>) (THOMAS, SAITO & SÆTROM, 2011).

Cada um deles, utilizando diferentes algoritmos, produz resultados distintos de possíveis MirSNPs: o PolymiRTS Database cataloga 358.874 MirSNPs; o miRSNP, 414.510; e o miRSNPscore, 396.850. Esses dados foram agrupados e, utilizando uma estatística Bayesiana ingênua (uma formula matemática), com parâmetros estabelecidos por Von Mering *et al.* (2004), estabeleceu-se um escore padronizado para listar os MirSNPs em função da sua probabilidade de ocorrência. Quanto mais próximo de 1, maiores as chances da predição ser real. Assim, foram considerados os MirSNPs com escore acima de 0,7, chegando em um número de 338.894 MirSNPs.

#### 4.1.2 INTEGRAÇÃO COM DADOS DE GWAS

Os estudos genômicos de associação (GWAS, do inglês *genome wide association studies*) utilizam amostras de pacientes e de indivíduos controle para avaliar quais SNPs podem estar associados a alguma doença, ou seja, se algum dos alelos é encontrado em um dos grupos em uma frequência estatisticamente significativa.

Assim, os 338.894 MirSNPs, obtidos até então pelos participantes do projeto, foram confrontados com GWAS de câncer de mama, disponíveis no banco de dados GWAS Catalog (<https://www.ebi.ac.uk/gwas/>). Também foram utilizados os dados do Projeto dos 1000 Genomas e estabelecidos o desequilíbrio de ligação dos MirSNPs,



sendo selecionados aqueles com coeficiente  $r^2 > 0,9$  e  $D' = 1$ . Com esse processo de filtragem adicional foram selecionados 43 MirSNPs.

#### 4.1.3 INTEGRAÇÃO DE DADOS DE eQTL

Pelo fato de que vários SNPs estão localizados em regiões não-codificantes do genoma, torna-se menos óbvia a relação de influência destas variantes genéticas com o desenvolvimento de uma doença. No entanto, muitas dessas regiões não-codificantes são regiões regulatórias e mudanças nelas podem alterar o padrão de expressão de um *locus*. Assim, utilizando os MirSNPs selecionados verificou-se quais deles poderiam contribuir na alteração da expressão do gene em que estão inseridos.

Para isso foram usados bancos de dados com estudos sobre expressão gênica buscando descrever eQTL (do inglês *expression quantitative trait loci*). Em primeiro, o BloodeQTL browser (<http://genenetwork.nl/bloodeqtlbrowser/>) (WESTRA *et al.*, 2013) que utilizou em suas análises, mais de 5000 amostras diferentes de sangue periférico. Em segundo, o GTEx eQTL browser (<http://www.gtexportal.org/home/>) (LONSDALE *et al.*, 2013), que adquiriu amostras coletadas de vários tecidos, incluindo tecido mamário, genotipou o material e analisou a expressão no nível de RNA. Por último, o METABRIC (CURTIS *et al.*, 2012), que contém os dados do genoma e do transcriptoma de 2000 amostras de tumores de mama.

Com a aplicação desses cis-eQTL, uma segunda filtragem para MirSNPs com um efeito eQTL resultou 17 MirSNPs.

Assim, foram selecionados MirSNPs com maiores chances de estarem no sítio alvo de um miRNA, apresentarem um efeito eQTL sobre seu gene e estarem associados com câncer de mama através dos dados fornecidos nos GWAS. Neste estudo, foram escolhidos os SNPs: rs8028; rs13005388; e rs71350312.

#### 4.2 LINHAGENS UTILIZADAS

Para que fossem estudadas as variações genotípicas em função dos rs8028, rs13005377 e rs71321208, utilizou-se o material genético de oito linhagens celulares imortalizadas. As linhagens celulares que foram utilizadas neste projeto foram

cedidas gentilmente pelo Laboratório de Citogenética Humana e Oncogenética (LabCHO) do Departamento de Genética da UFPR. Cada linhagem apresenta características próprias, mas todas foram obtidas a partir de um paciente do sexo feminino e são derivadas do tecido epitelial (QUADRO1).

**QUADRO 1. CARACTERÍSTICAS DAS LINHAGENS DE MAMA.**

Linhagens	classificação	Receptor Expresso	Doença	Idade	Origem
MCF-7	Luminal A	ER+; PR+/-; HER2-	Adenocarcinoma	69 anos	Caucasiana
ZR-75-1	Luminal B	ER+; PR+/-; HER2+	Carcinoma Ductal Infiltrado	63 anos	Caucasiana
BT474	Luminal B	ER+; PR+/-; HER2+	Carcinoma Ductal Invasivo	60 anos	Caucasiana
MDA-MB 468	Basal-like	ER-, PR-, HER2-	Adenocarcinoma Metastático	59 Anos	Afrodescendente
SKBR3	HER2	ER-, PR-, HER2+	Adenocarcinoma	43 anos	Caucasiana
MDA-MB 231	Triplo negativo	ER-, PR-, HER2-	Adenocarcinoma	51 anos	Caucasiana
MDA-MB 436	Triplo negativo	ER-, PR-, HER2-	Carcinoma Ductal Invasivo	43 anos	Caucasiana
184B5	Normal	ER+, PR+, HER2+	Sem doença	21 anos	Não informado

FONTE: HOLLIDAY & SPEIRS, 2011; NEVE *et al.*, 2006; WALEN & STAMPFER, 1989; CAILLEAU *et al.*, 1978; ENGEL *et al.*, 1978; LASFARGUES *et al.*, 1978; FOGH, J.; FOGH, J. M.; ORFEO, T, 1977; CAILLEAU *et al.*, 1974; SOULE *et al.*, 1973.

As linhagens celulares foram cultivadas em estufa com 5% de CO<sub>2</sub> a 37°C em meio de cultura RPMI 1640 (Gibco™) suplementado com 10% soro fetal bovino. Após a expansão das células, e confluência de 70-90%, as células foram tratadas com TrypLE™ Express (Gibco) (tripsina 0,25%) para desagregação das células do frasco de cultura. As células foram então submetidas à extração de DNA total seguindo o protocolo do All prep DNA mini kit (QIAGEN). O DNA foi utilizado como molde na amplificação da 3'UTR de *STAM2* e *LMOD1*.

#### 4.3 GENOTIPAGEM

Para verificar os alelos dos SNPs rs8028 de *LMOD1* e rs13005388 e rs71350312 de *STAM2* nas linhagens celulares de mama, foi realizado o

sequenciamento da porção de 3'UTR de cada gene, a partir do DNA previamente extraído. A amplificação da região de interesse foi feita através de reação em cadeia da polimerase (PCR) usando DNA em diferentes concentrações iniciais dependendo da linhagem. Também foi utilizado o tampão 1X de PCR (Invitrogen), 2mM de MgCl<sub>2</sub>, 0,2 mM de dNTP, 1 U Taq Platinum I (Invitrogen) e 10 pmoles de iniciadores específicos, em um total de 25 µl. Os iniciadores específicos utilizados para a amplificação da região 3'UTR contendo os rs8028, rs13005388 ou rs71350312 foram:

**QUADRO 2.** INICIADORES ESPECÍFICOS PARA REGIÃO CONTENDO O SNP DE INTERESSE.

SNP	Iniciador Direto	Iniciador Reverso
rs8028	CACTTGTAGGTTTCTCTGCCT	GAAGGAAGGGGAGCTGTTAAT
rs13005388	TCCCCAGAGGCCAGATACAGT	GCATCCACTGGGGATCTTGC
rs71350312	CCTCTCTTATTCTACAAGGCCA	TGTTTGGGGGACTGACTAAA

FONTE: O Autor (2017).

As condições de amplificação no termociclador foram:

**QUADRO 3 .**CICLOS DA PCR PARA AMPLIFICAÇÃO DA REGIÃO CONTENDO O SNP DE INTERESSE.

Etapas do Ciclo	Temperatura	Tempo	Número de Ciclos
Desnaturação Inicial	94°C	5 minutos	1
Desnaturação	94°C	30 segundos	35
Anelamento	57°C	30 segundos	
Extensão final	72°C	1 minuto	
Espera}	4°C	Indeterminado	

FONTE: O Autor (2017).

O gel de agarose a 1% serviu para averiguar se a PCR realmente amplificou o produto de interesse, sendo usado como marcador o Fast Gene50pb (Nippon Genetics). Neste gel, a corrida de eletroforese ocorreu à 90 Volts por 45 minutos.

Em seguida, 5  $\mu\text{L}$  do produto amplificado foi purificado, realizando-se uma purificação enzimática usando 0,2 $\mu\text{L}$  de Exonuclease I (20 U/ $\mu\text{L}$ ), responsável por degradar os oligonucleotídeos de DNA de fita simples, e Fosfatase Alcalina Termosensitiva (FastAp) (20 U/ $\mu\text{L}$ ), responsável por retirar os grupos fosfatos das extremidades 3' e 5' do DNA, por amostra, com o intuito de degradar os reagentes da PCR. As condições para a purificação foram

**QUADRO 4. CICLOS DA PURIFICAÇÃO ENZIMÁTICA DOS PRODUTOS AMPLIFICADOS POR PCR**

Etapa do Ciclo	Temperatura	Tempo	Número de Ciclos
Ação enzimática	37°C	60 minutos	1
Inativação da enzima	80°C	20 minutos	1
Em Espera	4°C	-	-

FONTE: O Autor (2017).

Após a purificação, as amostras foram, então, sequenciadas pelo método de Sanger. Foram usados 1  $\mu\text{L}$  de amostra purificada, 1,9  $\mu\text{L}$  de tampão, 0,5  $\mu\text{L}$  de BigDye e 0,1 $\mu\text{L}$  de iniciador específico para cada SNP. Para o SNP rs 8028 foi usado o iniciador reverso, e para o rs 71530312 o iniciador direto. As condições de sequenciamento foram:

**QUADRO 5. ETAPAS DO CICLO DE SEQUENCIAMENTO DAS AMOSTRAS PURIFICADAS**

Etapas do Ciclo	Temperatura	Tempo	Número de Ciclos
Desnaturação Inicial	95°C	1 minuto	1
Desnaturação	95°C	1 minuto	25
Anelamento	50°C	5 segundos	
Extensão	60°C	4 minutos	
Em Espera	4°C	1 minuto	-

FONTE: O Autor (2017).

Em seguida, as amostras foram purificadas por precipitação usando 2,5  $\mu\text{L}$  de ácido etilenodiamino tetra-acético (EDTA) a 125 Mm e 30  $\mu\text{L}$  de etanol absoluto por amostra. A placa contendo as amostras foi invertida sucessivamente e deixada a temperatura ambiente por 15 minutos. Então, a placa foi centrifugada por 45 minutos a 18°C e 3220 g; o sobrenadante foi descartado e 30  $\mu\text{L}$  de etanol 70% foram

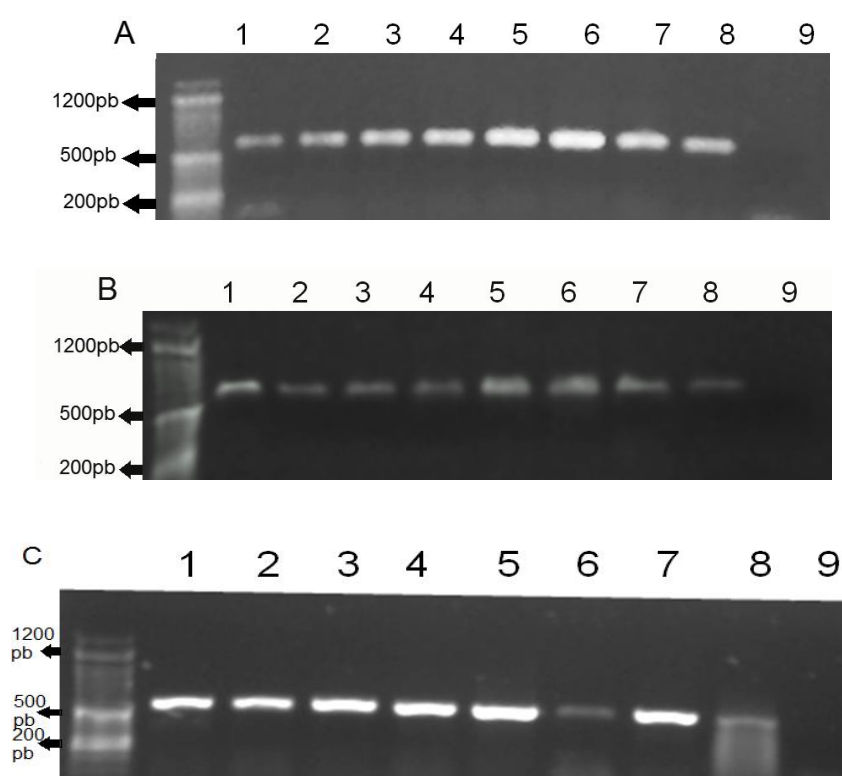
adicionados por amostra. A placa foi centrifugada por 5 minutos a 25°C, 3220 g, o sobrenadante foi descartado e a placa foi deixada por 20 minutos em temperatura ambiente para a evaporação do etanol restante. As amostras foram ressuspensas com 10µL de Hi-Di Formamida e levadas para o sequenciador (3500xl Genetic Analyser da Life Technologies). Os resultados foram analisados com o software Mutation Surveyor (SoftGenetics®).

## 5 RESULTADOS

### 5.1 AMPLIFICAÇÃO DA REGIÃO 3'UTR DO GENES *STAM2* E *LMOD1*

Para saber se as predições realizadas pela ferramenta de bioinformática realmente ocorrem, inicialmente foi feita a amplificação de parte da região 3' UTR dos genes *STAM2* e *LMOD1*, nas quais se encontram os SNPs de interesse, utilizando o DNA das linhagens tumorais e normal de mama. O produto da amplificação do rs8028, rs71350312 e do rs13005388 apresentaram tamanhos correspondentes a 600, 627e 509 pares de base respectivamente. Os resultados da reação de PCR podem ser visualizados na FIGURA 5:

**FIGURA 5:** ELETROFORESE DA PCR DA REGIÃO 3'UTR DOS GENES *STAM2* E *LMOD1*

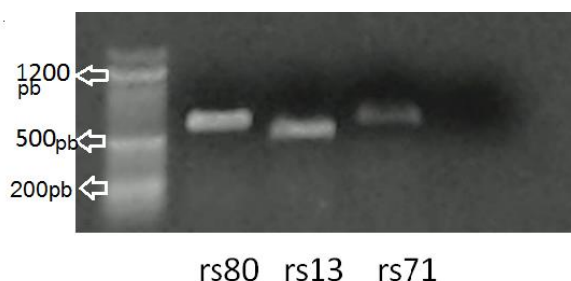


Na figura 4 é possível ver os resultados da PCR. Em A temos a PCR do rs8028. Em B a PCR do rs71350312. Em C a PCR do rs13005388. 1- Controle Positivo, 2-SKBR3, 3-BT 474, 4-MCF-7, 5-ZR 75-1, 6-184B5, 7-MDA MB 468, 8-MDA MB 436 e 9-BRANCO (controle negativo da reação). Fonte: O autor (2017).

Para a amostra MDA-MB-231 foi feito um gel de eletroforese separado, pois os primers foram testados com o DNA de uma das linhagens, para avaliar se

funcionavam, antes de serem usados nas demais linhagens, que pode ser visualizado na FIGURA 6.

**FIGURA 6:** GEL DE ELETROFORESE DA LINHAGEM MDA-MB-231



Na FIGURA 5 é possível observar a amplificação da região 3' UTR, da linhagem MDA-MB-231. Em rs80: amplificação da região que contém o rs8028, do gene *LMOD1*. Em rs13: amplificação da região que contém o rs13005388, do gene *STAM2*. Em rs71 que contém o rs71350312, do gene *STAM2*. Fonte: O autor (2017).

## 5.2 SEQUENCIAMENTO

Após a verificação de que a reação de PCR ocorreu corretamente, as amostras foram sequenciadas pelo Método de Sanger, e o resultado, com o genótipo de cada amostra pode ser visto na TABELA 6.

**QUADRO 6.** GENÓTIPO DAS LINHAGENS DE MAMA

Linhagens de Mama	rs71350312	rs13005388	rs8028
	Genótipo	Genótipo	Genótipo
SKBR3	T / T	A / A	T / C
ZR-75-1	T / C	A / A	C / C
BT-474	T / T	A / A	T / T
184-B5	T / T	A / A	C / C
MCF-7	T / T	A / A	C / C
MDA-MB-231	Inconclusivo	Inconclusivo	C / C
MDA-MB-468	T / T	A / A	T / C
MDA-MB-436	C / C	A / G	T / T

Fonte: O Autor (2017).

Dois dos SNP (rs71350312 e rs8028) estudados apresentam, cada um, dois alelos, sendo que em ambos esses alelos vão ser uma timina (T) e o outro uma citosina (C). Assim, como cada célula tem dois alelos, o genótipo delas pode ser homozigoto para timina, homozigoto para citosina, ou heterozigoto. O rs13005388 apresenta dois alelos a adenina (A) e a guanina (G), sendo que o genótipo das linhagens pode ser: homozigoto para adenina, homozigoto para guanina, ou heterozigoto.

Das linhagens genotipadas para o rs71350312, a maioria se mostrou homozigota para T, com exceção da linhagem ZR-75-1, que se mostrou heterozigota, e da linhagem MDA-MB-436 que é homozigota para o alelo C.

Já o rs13005388 apresentou quase todas as linhagens homozigotas para o alelo A, com exceção da linhagem MDA-MB-436 que apresentou um genótipo heterozigoto.

Para o rs8028 a maioria se mostrou homozigoto para C, as linhagens SKBR3 e MDA-MB-468 apresentaram um genótipo heterozigoto e as linhagens MDA-MB-436 e BT474 homozigotas para T.

Na análise de reação de sequenciamento do rs71350312 e do rs13005388, no software Mutation Surveyor (SoftGenetics®), a sequência da amostra MDA-MB-231 apresentou muitas diferenças com a sequência estudada, tornando seu resultado duvidoso para ser colocado nesse trabalho.

### 5.3. INTERAÇÃO ENTRE MIRNAS PREDITOS E O SNP

A predição realizada pela ferramenta de bioinformática sobre a existência de possíveis MirSNPs, relacionados ao câncer de mama, selecionou um conjunto de miRNAs, para cada um dos alelos do SNP associado. Para cada MirSNP o programa atribuiu um escore que vai de zero a 1, quanto mais próximo a 1 maior é a probabilidade da interação entre o miRNA e do SNP ocorrer (TABELA 7).

Nesse trabalho, foi utilizado um software *Target Scan Human*, que também realiza predições sobre interações entre miRNA e SNP alvos. Para isso, ele analisa a sequência da região *seed* e baseando-se na complementaridade das sequências, informa em quais sítios da 3'UTR do gene selecionado, nesse caso o *LMD1* e o *STAM2*, o miRNA pode se ligar.



**QUADRO 7. PREDIÇÕES REALIZADAS PARA OS MIRSNS**

SNP	Alelos	miRNA	Escore
rs71350312	C	hsa-miR-191-3p	0.8313787
	T	hsa-miR-508-5p	0.7606811
		hsa-miR-4793-3p	0.9317711
rs13005388	G	hsa-miR-6730-3p	0.9779717
		hsa-miR-877-3p	0.8559046
		hsa-miR-2682-3p	0.8044826
	A	hsa-miR-514a-3p	0.9651534
		hsa-miR-514b-3p	0.9651534
rs8028	C	hsa-miR-1296-3p	0.7079350
		hsa-miR-4283	0.7488324
		hsa-miR-4656	0.8467698
		hsa-miR-4489	0.7565631
		hsa-miR-194-3p	0.9075926
		hsa-miR-4740-5p	0.7552043
		hsa-miR-4648	0.9848377
		hsa-miR-185-3p	0.9882725
		hsa-miR-3151-5p	0.8592348
	T	hsa-miR-4740-5p	0.7552043
		hsa-miR-4648	0.9848377
		hsa-miR-3199	0.9920605
		hsa-miR-4513	0.9723837
		hsa-miR-8052	0.9925073
		hsa-miR-3660	0.7020974

Fonte: O autor (2017).

O *Target Scan Human* atribui uma pontuação para todas as interações preditas, elas apresentam um valor negativo, e quanto mais negativo for, maiores as

chances delas ocorrerem. A pontuação é gerada a partir de características como: o tipo de sítio (regular ou não-regular), se existe pareamento fora da região *seed*, quantidade de bases adenina e uracila no 3'UTR, distancia mínima de 15 nucleotídeos do códon de término, comprimento do 3'UTR, estabilidade entre o pareamento da região *seed* e do sítio-alvo (Agarwal et al.,2015) (TABELA 8).

Como é atribuída uma pontuação para cada interação predita, caso um mesmo miRNA se ligue em diferentes lugares na 3'UTR, vai existir para cada sítio-alvo um valor. Esses valores vão ser somados e o número apresentado na coluna 'escore' da TABELA 8 representa essa soma.

Esse software quando avaliou a interação dos miRSNPs preditos pela ferramenta de bioinformática revelou que muitos miRNAs conseguiam se ligar em diferentes sítios-alvos, além daquele que sítio que continha o SNP. O hsa-miR-4740-5p, hsa-miR-3199, hsa-miR-4513 e hsa-miR-8052, do gene *LMOD1* foram os únicos que apresentaram uma interação predita somente ao sítio-alvo do SNP. Os miRNAs dos outros dois SNPs ou apresentaram mais de um sitio alvo de ligação da região 3'UTR, ou não apresentaram ligação nenhuma com o alvo.

QUADRO 8. TARGET SCAN HUMAN

SNP	MicroRNAs	Escore	Números de sítios-alvos	Se liga no sítio alvo do SNP
rs71350312	hsa-miR-191-3p	Não tem		
	hsa-miR-4793-3p*	-0,15 ( <b>-0,02</b> / -0,03/ -0,11)	3	Sim
	hsa-miR-508-5p*	-0,18 ( <b>-0,02</b> / -0,16)	2	Sim
rs13005388	hsa-miR-6730-3p	Não tem		
	hsa-miR-877-3p	Não tem		
	hsa-miR-2682-3p	-0,28 (-0,06/ <b>-0,22</b> )	2	Sim
	hsa-miR-514a-3p	-0,26 (-0,05/ <b>-0,21</b> )	2	Sim
	hsa-miR-514b-3p	-0,26 (-0,05/ <b>-0,21</b> )	2	Sim
rs8028	hsa-miR-1296-3p*	-0,11	1	Não
	hsa-miR-4283*	-0,07	1	Não
	hsa-miR-4656*	-0,2	1	Não
	hsa-miR-4489*	-0,17	1	Não
	hsa-miR-194-3p*	-0,37 ( <b>-0,09</b> / -0,13/ -0,10)	3	Sim
	hsa-miR-4740-5p	-0,09	1	Sim
	hsa-miR-4648*	-0,51 ( <b>-0,17</b> / -0,34)	2	Não
	hsa-miR-185-3p*	-0,29	1	Não
	hsa-miR-3151-5p*	-0,39 (-0,26/ -0,13)	2	Não
	hsa-miR-3199	<b>-0,24</b>	1	Sim
	hsa-miR-4513	<b>-0,12</b>	1	Sim
	hsa-miR-8052	<b>-0,27</b>	1	Sim
	hsa-miR-3660*	-0,11 ( <b>-0,05</b> / -0,06)	2	Sim

Em asterisco miRNAs que possuem outro sítio-alvo na 3'UTR do gene alvo, no qual o SNP de estudo não está inserido. Para os miRNAs que possuem mais de um sítio alvo, em negrito está destacado o valor que o software atribuiu para a interação entre o miRNA e o sítio alvo, no qual o SNP é encontrado. Fonte: O Autor (2017).

## 6. DISCUSSÃO

Variações genéticas são encontradas ao longo do genoma de todos os indivíduos, sendo elas responsáveis pelas diferenças que normalmente são vistas em uma população. Essas mudanças muitas vezes estão relacionadas a patologias, podendo ter efeitos mais visíveis, no qual apenas a mudança de um nucleotídeo pode ser responsável pelo aparecimento da doença, como em doenças monogênicas, ou efeitos menos perceptíveis, como no caso do câncer onde é necessário ocorrerem diversas mutações, para que a doença manifeste-se (LEVY *et al.*, 2007).

Entre as variações genéticas que existem, a mais comum é o polimorfismo de nucleotídeo único (do inglês SNP, *single nucleotide polymorphism*) (FURRER *et al.*, 2016). Procurando analisar como essas variações influenciavam no aparecimento do câncer, foram realizados muitos estudos com SNP que estavam em regiões codificantes, nas quais a mutação de um par de bases no DNA podia alterar a sequência de aminoácidos, gerando uma molécula defeituosa. Contudo, muitos desses polimorfismos são silenciosos, não alterando a sequência de aminoácidos, e por essa razão começou a se investigar SNP que estavam em regiões regulatórias e como essas variações poderiam causar modificações na expressão do gene em que são encontradas.

A região regulatória de interesse nesse trabalho é a região 3'UTR, que é encontrada em mRNA maduros, sendo que nela estão localizados os sítios alvos para os miRNAs. Os miRNAs são moléculas capazes de modificar o padrão de expressão gênica nas células, realizando uma regulação pós-transcricional. Esses RNAs são muito pequenos, entre 19 a 22 nucleotídeos, e dentro dessa sequência existe uma região altamente conservada chamada *seed*, encontrada na extremidade 5' sendo formada a partir do segundo até o oitavo nucleotídeo. O pareamento da região *seed* com o sítio alvo é o principal responsável pela ligação entre o miRNA e o mRNA. Assim, a mudança de apenas um par de base no sítio alvo pode: inibir a interação entre as duas moléculas, fortalecer a ligação do miRNA, ou criar um novo sítio alvo para outro miRNA (BARTEL, 2004).

Estudos já mostraram que SNPs no sítio alvo, pode interferir na regulação gênica, aumentando o risco a certas doenças. Tchatchou e colaboradores em 2009 analisaram 11 miRSNPs que foram relacionados ao câncer de mama. Fazendo um

estudo de caso-controle em pacientes com alto risco de desenvolver câncer de mama familiar, constataram que apenas um miRSNP apresentava uma correlação positiva com esses pacientes, que nesse caso envolveu o alelo variante do rs2747648, que fica localizado no sítio alvo do miR-453, na 3'UTR gene *ESR1*. O gene *ESR1* codifica para o receptor de estrogênio alfa, e o rs2747648 apresenta os alelos T e C, sendo que na presença do alelo C o miR-453 consegue interagir com o mRNA desse gene, reduzindo sua expressão, porém em genes com o alelo T esse evento não acontece. Isso acaba promovendo um aumento da expressão desse gene, e consequentemente no número de receptores de estrogênio alfa, uma característica de aumenta o risco ao câncer de mama, principalmente dos tipos ER+.

Além desse, outros estudos já mostraram como miRNAs podem ter um importante efeito sobre linhagens tumorais, tanto que são classificados como oncogenes ou genes supressores de tumor. Guttilla e White em 2009 apontaram que o miR-27a, o miR-96 e o miR-182, tinham como alvo em comum o mRNA do gene *FOXO1*, supressor tumoral, e que na linhagem de tumoral de mama, MCF-7, tais miRNAs estavam com a expressão aumentada. Assim, conseguindo promover a inibição deles, foi constatado uma diminuição do crescimento dessa linhagem e um aumento na apoptose dessas células.

Outro estudo mostrou ainda como miRNAs podem auxiliar no tratamento contra o câncer, por exemplo, em linhagens celulares resistentes à cisplatina. A cisplatina é um quimioterápico capaz de promover quebras na fita de DNA levando as células cancerígenas à morte. HE *et al* em 2015, mostrou que linhagens MCF-7 resistentes à cisplatina, tinham baixas concentrações de miR-218 quando comparados com a linhagem não resistente. Ao realizar o ensaio de transfecção de mimetizadores de miR-218 nas células resistentes, percebeu uma diminuição do crescimento delas, e um aumento no processo de apoptose. Assim, analisou possíveis alvos desse miRNA e constatou que o gene *BRCA1* era um de seus alvos. Uma vez que a função desse gene é realizar reparo no DNA, foi discutido no trabalho que talvez o aumento do miR-218, diminui a presença do mRNA de *BRCA1*, permitindo a droga apresentar um efeito anti-neoplásico mais evidente em células que inicialmente eram resistentes a ela.

Com isso, é possível perceber que os miRNAs apresentam diferentes efeitos no metabolismo de células tumorais dependendo do gene em que atuam. Assim, a

presença de variações genética nos sítios alvos de miRNAs podem modificar o padrão da expressão gênica dessas células, dependendo do gene e do tipo de células em que eles ocorrem, o que pode favorecer, ou não, o aparecimento da doença. Contudo, essas variações só vão conseguir gerar alguma mudança no padrão de regulação das células, se os genes estiverem sendo expressos nelas.

Nesse trabalho foi avaliado a variação genética que ocorre na 3'UTR dos genes *LMOD1* e *STAM2* em oito diferentes linhagens de mama. Consultando o *The Human Protein Atlas*, que utilizou 214 amostras de tecido mamário, foi observado a expressão desses genes em tecido mamário. O gene *LMOD1* apresentou um valor médio de 68,1 RPKM (Leituras por Kilobase por Milhão, do inglês *Reads Per Kilobase Per Million*), enquanto o *STAM2* teve um valor médio de 6,4 RPKM. O Atlas também informou sobre a presença da proteína codificada por esses genes. Experimentos utilizando anticorpos monoclonais em 3 diferentes amostras, revelou que tanto a proteína *LMOD1*, quanto a *STAM2* apresentam um padrão de expressão média no tecido mamário. Portanto a presença de SNP na região 3'UTR, podem interferir na regulação desses dois genes, e consequentemente na produção de mRNA e proteínas.

As informações geradas pelas ferramentas de bioinformática selecionaram três SNP que estavam associados ao câncer de mama e apresentavam um efeito cis-eQTL, o rs8028 do *LMOD1* e o rs71350312 e o rs13005388 *STAM2*, sendo que para cada um deles foi selecionado um conjunto de miRNAs que podem interagir com os diferentes alelos dos SNP (TABELA 7).

O propósito desse trabalho foi determinar a existência de variação genética nas linhagens de mama disponíveis para o estudo (TABELA 5), nos SNPs preditos. Caso todas as linhagens apresentassem o mesmo genótipo não seria possível ver como os diferentes alelos dos SNP regulam o gene em que estão inseridos, ou a realização de ensaios funcionais que validassem as predições de miRSNPs criadas pela ferramenta de bioinformática. Como essas variações genéticas foram encontradas, isso permite a continuação de mais estudos envolvendo os SNPs, e os miRNAs relacionados.

## 7. CONCLUSÃO

Com a realização da reação de sequenciamento pelo Método de Sanger, foi possível determinar os genótipos das linhagens para o SNP de interesse, e assim comprovar a existência de uma variação genética entre as linhagens. Para o rs8028, duas linhagens apresentaram o genótipo heterozigoto MDA-MB-468 e SKBR3, duas foram homozigotas para T, BT-474 e MDA-MB-436 e as quatro restantes foram homozigotas para C, MCF-7, ZR-75-1, 184-B5 e MDA-MB-231. Para o rs71350312 apenas uma linhagem é heterozigota a ZR-75-1, entre as linhagens homozigotas, apenas uma também, a MDA-MB-436 mostrou duas cópias do alelo C, enquanto as outras 5 restantes são homozigotas para T. Para linhagem MDA-MB-231 não foi possível determinar o genótipo devido à incompatibilidade entre a sequência obtida e a sequência padrão. Para o rs13005388 quase todas as linhagens apresentaram um genótipo homozigoto para o alelo A, enquanto apenas a linhagem MDA-MB-436 mostrou-se heterozigota, e para a linhagem MDA-MB-231 também não foi possível determinar o genótipo dessa linhagem.

## 8. PERSPECTIVA

Para que as interações preditas entre os miRSNPs sejam confirmadas é necessário a realização de um ensaio funcional, que não foi iniciado ainda. O experimento mais utilizado para obter essa confirmação é o ensaio com o gene reporter luciferase, no qual é necessário o DNA de duas linhagens, cada uma homozigota para um dos alelos, e a escolha de um miRNA teste, que vai interagir com, um dos alelos, porém não com o outro. Nesse trabalho ainda foi realizado uma pesquisa, utilizando o software *Target Scan* para selecionar qual dos miRNA preditos seria de melhor interesse para realizar os ensaios funcionais.

Para o rs8028 temos os miRNAs hsa-miR-3199 e hsa-miR-8052 que apresentaram os melhores escores pela predição desenvolvida pelo nosso grupo e pelo Target Scan: 0.9920605, -0,24 e 0.9925073, -0,27 respectivamente. Além disso de acordo com o Target Scan esse dois miRNAs se ligam em apenas um sitio na 3'UTR do *LMOD1*, que é o sitio onde o SNP está inserido.

Para o rs71350312, a ferramenta de predição informou apenas 3 prováveis miRNAs, sendo que dois deles apresentaram, pelo Target Scan, mais de um sítio alvo. Assim apesar do Target Scan não relacionar o hsa-miR-191-3p com nenhum sítio alvo na 3'UTR do *STAM2*, esse miRNA se torna uma escolha interessante pelo fato do seu escore, de acordo com as predições realizadas pelo grupo dele se ligar ao alelo C do SNP e possuir um escore de 0,83.

Para o rs13005388 dos 5 miRNAs preditos dois deles pelo Target Scan não conseguiam interagir com o sítio-alvo, e os três restantes (hsa-miR-2682-3p, hsa-miR-514a-3p e o hsa-miR-514b-3p) apresentaram dois possíveis sítios alvos, sendo que em um deles o SNP esta localizado. Nos três casos esse sitio, que contem o SNP, apresentou o maior escore. Assim considerando que as ferramentas de bioinformática atribuiu um escore de 0,96 ao miR-514-3p, esse seria a melhor escolha para a realização de futuros ensaios funcionais.



## 9 REFERÊNCIAS

AGARWAL, V. et al. Predicting effective microRNA target sites in mammalian mRNAs. **Elife**, v. 4, 2015.

AMERICAN SOCIETY OF CANCER. Disponível em: <https://www.cancer.org/cancer/breast-cancer/about/what-is-breast-cancer.html>. Acesso em 29/03/2017.

BACHE, K. G.; RAIBORG, C.; MEHLUM, A.; STENMARK, H. STAM and Hrs Are Subunits of a Multivalent Ubiquitin-binding Complex on Early Endosomes. **THE JOURNAL OF BIOLOGICAL CHEMISTRY**, v. 278, n. 14, p. 12513–12521, 2003.

BARTEL, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. **Cell**, v. 116, n. 2, p. 28–97, 2004.

BECKETTA, E. L.; VEYSEYB, M.; LUCOCK, M. Folate and microRNA: Bidirectional interactions. **Clinica Chimica Acta**, v. 474, p. 60-66, 2017.

BENITEZ, J. et al. Advances in Brief A Region of Allelic Imbalance in 1q31-32 in Primary Breast Cancer Coincides with a Recombination Hot Spot. **CANCER RESEARCH**, v.57, p, 4217-4220, 1997.

BERINDAN-NEAGOE, I. et al. MicroRNAome genome: a treasure for cancer diagnosis and therapy. **CA Cancer J Clin**, v. 64, n. 5, p. 311–336, 2014.

BHATTACHARYA, A.; ZIEBARTH, J.D.; CUI, Y. PolymiRTS Database 3.0: linking polymorphisms in microRNAs and their target sites with human diseases and biological pathways. **Nucleic Acids Research**, v. 42 Database issue, p. 86- 91, 2014.

BOCZKOWSKA, M. et al. How Leiomodin and Tropomodulin use a common fold for different actin assembly functions. **Nature Communications**, v. 6, n. 8314, 2015.

BOHON, C. Cancer Recognition and Screening for Common Breast Disorders and Malignancy. **Obstetrics and Gynecology Clinics of North America**, v. 44, n. 2, p. 257 - 27, 2017.

BORCHERT, G. M.; LANIER, W.; DAVIDSON, B. L. RNA polymerase III transcribes human microRNAs. **Nat Struct Mol Biol**, v. 13, n. 12, p. 1097–1101, 2006.

CAILLEAU, R. et al. Long-term human breast carcinoma cell lines of metastatic origin: preliminary characterization. **In Vitro**, v. 14, p. 911-915, 1978.

CAILLEAU R, et al. Breast tumor cell lines from pleural effusions. **J. Natl. Cancer Inst**, v. 53, n. 3, p. 661-674, 1974

CHRYSANTHOU, E. et al. Phenotypic characterisation of breast cancer: the role of CDC42. **Breast Cancer Res Treat**, p. 1-9, 2017.

CIRQUEIRA, M. B. et al. Subtipos moleculares do câncer de mama. **Femina**, v. 39, n. 10, 2011.

CONLEY, C. A. et al. Leiomodins: Larger Members of the Tropomodulin (Tmod) Gene Family. **Genomics**, v. 73, p. 127–139, 2001.

CONLEY, C. A. Leiomodins and tropomodulin in smooth muscle. **Am J Physiol Cell Physiol**, v. 280, p. 1645–1656, 2001.

COWPER-SALLARI, R. et al. Breast cancer risk-associated SNPs modulate the affinity of chromatin for FOXA1 and alter gene expression. **Nature Genetics**, v. 44, n. 11, p. 1191–1198, 2012.

CURTIS, C. et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. **Nature**, v. 486, n. 7403, p. 346–352, 2012.

DLLINK. Disponível em: <https://analysistools.nci.nih.gov/LDlink/>. Acessado em: 21/06/2017.

ELKIN, S. R.; LAKODUK, A. M.; SCHMID, S. L. Endocytic Pathways and Endosomal Trafficking: A Primer. **Wien Med Wochenschr**, v. 166, n. 7-8, p. 196–204, 2016.

ENDO, K. et al. STAM2, a new member of the STAM family, binding to the Janus kinases. **FEBS Letters**, v. 477, n. 1-2, p. 55–61, 2000.

ENGEL, L. W. et al. Establishment and Characterization of Three New Continuous Cell Lines Derived from Human Breast Carcinomas. **CANCER RESEARCH**, v. 38, p. 3352 - 3364, 1978.

ENSEMBL. Disponível em: <http://www.ensembl.org>. Acessado em: 06/06/2017.

FANG, L. L. et al. Expression, regulation and mechanism of action of the miR-17-92 cluster in tumor cells. **INTERNATIONAL JOURNAL OF MOLECULAR MEDICINE**, v. 40, p. 1624–1630, 2017.

FERLAY, J.; SOERJOMATARAM, I.; DIKSHIT, R.; ESER, S.; MATHERS, C.; REBELO, M.; PARKIN, D. M.; FORMAN, D.; BRAY, F. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. **International Journal of Cancer**, v. 136, p. 359–386, 2014.

FOGH, J.; FOGH, J. M.; ORFEO, T. One hundred and twenty-seven cultured human tumor cell lines producing tumors in nude mice. **Journal of the National Cancer Institute**, v. 59, p. 221–26, 1977.

FRAZER, K. A. et al. Human genetic variation and its contribution to complex traits. **Nature Reviews**, v. 10, p. 241–251, 2009.

FURRER, D. et al. Evaluation of human epidermal growth factor receptor 2 (HER2) single nucleotide polymorphisms (SNPs) in normal and breast tumor tissues and their link with breast cancer prognostic factors. **The Breast**, v. 30, p. 191-196, 2016.

GENECARD. Disponível em: <http://www.genecards.org/>. Acesso em: 20/05/2017.

GUTTILLA, I.K.; WHITE, B. A. Coordinate Regulation of FOXO1 by miR-27a, miR-96, and miR-182 in Breast Cancer Cells. **THE JOURNAL OF BIOLOGICAL CHEMISTRY**, v. 284, n.35, p. 23204-23216, 2009.

HAJARNIS, S. et al. Suppression of microRNA Activity in Kidney Collecting Ducts Induces Partial Loss of Epithelial Phenotype and Renal Fibrosis. **Journal American Society Nephrology**, v. 29, 2017.

HALIM, D. et al. Loss of LMOD1 impairs smooth muscle cytocontractility and causes megacystis microcolon intestinal hypoperistalsis syndrome in humans and mice. **Proc Natl Acad Sci USA**, v. 114, n. 13, p. 2739-2747, 2017.

HANSON, P.I.; CASHIKAR, A. Multivesicular Body Morphogenesis. **Annu Rev Cell Dev Biol**, v. 28, p. 337-362, 2012.

HANSON, P. I.; SHIM, S. MERRILL, S. A. Cell biology of the ESCRT machinery. **Current Opinion in Cell Biology**, v. 21, p.568-574, 2009.

HANYALOGU, A. C.; MCCULLAGH, E.; VON ZASTROW, M. Essential role of Hrs in a recycling mechanism mediating functional resensitization of cell signaling. **EMBO J**, v. 24, n. 13, p. 2265-2283, 2005.

HE, X. et al. MiR-218 regulates cisplatin chemosensitivity in breast cancer by targeting BRCA1. **Tumor Biol**, v. 36, n. 3, p. 2065-2075, 2015.

HOLLIDAY, D. L.; SPEIRS, V. Choosing the right cell line for breast cancer research. **Breast cancer Res**, v. 13, n. 4, p. 1-7, 2011.

HURLEY, J. H. ESCRTs are everywhere. **EMBO J**, v. 34, n. 19, p. 2398-2407, 2015.

HURLEY, J. H. The ESCRT Complexes. **Crit Rev Biochem Mol Biol**, v. 45, n. 6, p. 463-487, 2010.

INCA. Instituto Nacional do Câncer. Disponível em: <http://www2.inca.gov.br>. Acesso em: 06/06/2017.

JARA, L. et al. Mutations in BRCA1, BRCA2 and other breast and ovarian cancer susceptibility genes in Central and South American populations. **Biological Research**, v. 50, n. 35, 2017.

- KAPURALIN, K. et al. STAM2, a member of the endosome-associated complex ESCRT-0 is highly expressed in neurons. **Molecular and Cellular Neuroscience**, v. 67, p. 104-115, 2015.
- KATZMANN, D. J.; BABST, M.; EMR, S.D. Ubiquitin-Dependent Sorting into the Multivesicular Body Pathway Requires the Function of a Conserved Endosomal Protein Sorting Complex, ESCRT-I. **Cell**. v.106, p. 145-155, 2001.
- KOMADA, M.; SORIANO, P. Hrs, a FYVE finger protein localized to early endosomes, is implicated in vesicular traffic and required for ventral folding morphogenesis. **Genes Dev**, v.13, n.11, p. 1475-1485, 1999.
- KOSTYUKOVA, A. S. Leiomodin/tropomyosin interactions are isoform specific. **Biochemistry and Biophysics**, v. 465, n. 1, p. 227-230, 2007.
- KOUTRAS, A. K.; EVANS, T. R. J. The epidermal growth factor receptor family in breast cancer. **Onco Targets and Ther**, v. 1, p. 5-19, 2008.
- LASFARGUES, E.Y. et al. Isolation of two human tumor epithelial cell lines from solid breast carcinomas. **J. Natl. Cancer Inst**, v. 61, p. 967-978, 1978.
- LEVY, S. et al. The Diploid Genome Sequence of an Individual Human. **PLoS BIOLOGY**, v. 5, n.10, p. 2113-2144, 2007.
- LIU, C. et al. MirSNP, a database of polymorphisms altering miRNA target sites, identifies miRNA-related SNPs in GWAS SNPs and eQTLs. **BMC Genomics**, v. 13, n. 661, p. 1-10, 2012.
- LONSDALE, J. et al. The Genotype-Tissue Expression (GTEx) project. **Nature Genetics**, v. 45, p. 580-585, 2013.
- MATTISSEK, C.; TEIS, D. The role of the endosomal sorting complexes required for transport (ESCRT) in tumorigenesis. **Mol Membr Biol**, v. 31, n. 4, p. 111-119, 2014.
- MIANO, J. M. Serum response factor: toggling between disparate programs of gene expression. **Journal of Molecular and Cellular Cardiology**, v. 35, p. 577–593, 2003.
- NAMI B.; WANG Z. HER2 in Breast Cancer Stemness: A Negative Feedback Loop towards Trastuzumab Resistance. **Cancers**, v.9, n.5, p.40-73, 2017.
- NANDA, V.; MIANO, J. Leiomodin 1, a New Serum Response Factor-dependent Target Gene Expressed Preferentially in Differentiated Smooth Muscle Cells. **The Journal of Biological Chemistry**, v. 287, n. 4, p. 2459 –2467, 2011.
- NATIONAL CANCER INSTITUTE. Disponível em: <https://www.cancer.gov/>. Acessado em 23/05/2017.
- NEVE, R. M. et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. **Cancer Cell**, v. 10, n. 6, p. 515–527, 2006.

PEROU, C. M. et al. Molecular portraits of human breast tumours. **Nature**, v.406, p. 747-752, 2000.

PRESKILL, C.; WEIDHAAS, J. B. SNPs in microRNA binding sites as prognostic and predictive cancer biomarkers. **Crit Rev Oncog**, v. 18, n. 4, p. 327-340, 2013.

RAKHA, E. A.; ELLIS I. O. Triple-negative/basal-like breast cancer: review. **Royal College of Pathologists of Australasia**, v. 41, n.1, p.40-47, 2009.

RAZI, M.; FUTTER, C. E. Distinct Roles for Tsg101 and Hrs in Multivesicular Body Formation and Inward Vesiculation. **Molecular Biology of the Cell**, v. 17, p. 3469–3483, 2006.

REN, X.; HURLEY, J. H. VHS domains of ESCRT-0 cooperate in high-avidity binding to polyubiquitinated cargo. **The EMBO Journal**, v. 29, p. 1045-1054, 2010.

RODGERS, K. M. ; UDESKY, J. O.; RUDEL, R. A.; BRODY, J. G. Environmental chemicals and breast cancer: An updated review of epidemiological literature informed by biological mechanisms. **Environmental Research**, v. 160, p.152-182, 2017.

SAKORAFAS, G. H. The management of women at high risk for the development of breast cancer: risk estimation and preventative strategies. **Cancer Treatment Reviews**, v. 29, p. 79–89, 2003.

SØRLIE, T. et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. **Proc Natl Acad Sci USA**, v. 98, n. 19, p. 10869-10874, 2001.

SOULE, H. D. et al. A human cell line from a pleural effusion derived from a breast carcinoma. **J. Natl. Cancer Inst**, v. 51, p. 1409-1416, 1973.

TAKATA, H. et al. A Hrs binding protein having a Src homology 3 domain is involved in intracellular degradation of growth factors and their receptors. **Genes to Cell**, v. 5, n. 1, p. 57-69, 2000.

TARGET SCAN HUMAN. Disponível em:< [http://www.targetscan.org/vert\\_71/](http://www.targetscan.org/vert_71/)> . Acessado em 24/ 11/2017.

TCHATCHOU, S. et al. A variant affecting a putative miRNA target site in estrogen receptor (ESR) 1 is associated with breast cancer risk in premenopausal women. **Carcinogenesis**, v. 30, n. 1, p. 59–64, 2009.

TELANG, N. T. et al. The nutritional herb *Epimedium grandiflorum* inhibits the growth in a model for the Luminal A molecular subtype of breast cancer. **Oncology Letters**, v.13, p. 2477-2482, 2017.

THE HUMAN ATLAS PROTEIN. Disponível em <<https://www.proteinatlas.org/>>. Acessado em 24/11/2017.

THOMAS LF, SAITO T, SÆTROM P. Inferring causative variants in microRNA target sites. **Nucleic Acids Research**, v. 39, n. 16, e109, p. 1-10, 2011.

VISONE, R.; CROCE, C. M. MiRNAs and Cancer. **The American Journal of Pathology**, v. 174, n.4, p. 1131–1138, 2009.

VON MERING, C. et al. STRING: known and predicted protein-protein associations, integrated and transferred across organisms. **Nucleic Acids Research**, v. 33, D433-437, 2005.

WALEN, K. H.; STAMPFER, M. R. Chromosome Analyses of Human Mammary Epithelial Cells at Stages of Chemical-Induced Transformation Progression to Immortality. **Cancer Genetics**, v 37, n. 2, p. 249-262, 1989.

WEIGEL, M. T.; DOWSETT, M. Current and emerging biomarkers in breast cancer: prognosis and prediction. **Endocr Relat Cancer**, v. 17, n. 4, p. 245-62, 2010.

WESTRA, H.J. et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. **Nature Genetics**, v. 45, p. 1238-1243, 2013.

YAMADA, M. et al. Signal-Transducing Adaptor Molecules STAM1 and STAM2 Are Required for T-Cell Development and Survival. **Molecular and Cellular Biology**, v. 22, n. 24, p. 8648-8658, 2002.

YUEN, M. *et al.* Leiomodin-3 dysfunction results in thin filament disorganization and nemaline myopathy. **The Journal of Clinical Investigation**, v. 124, n.11, p.4693-4407, 2015.

ZHOU, C. Z. et al. Refined mapping of loss of heterozygosity on 1q31.1-32.1 in sporadic colorectal carcinoma. **World J Gastroenterol**, v. 14, n. 10, p. 1582-1587, 2008.